Weyl variations and local sufficiency of linear observers in the mean square optimal coherent quantum filtering problem*

Igor G. Vladimirov†

In loving memory of my mother, Raisa Dmitrievna Vladimirova (10.05.1933 – 30.05.2015)

Abstract

This paper is concerned with the coherent quantum filtering (CQF) problem, where a quantum observer is cascaded in a measurement-free fashion with a linear quantum plant so as to minimize a mean square error of estimating the plant variables of interest. Both systems are governed by Markovian Hudson-Parthasarathy quantum stochastic differential equations driven by bosonic fields in vacuum state. These quantum dynamics are specified by the Hamiltonians and system-field coupling operators. We apply a recently proposed transverse Hamiltonian variational method to the development of first-order necessary conditions of optimality for the CQF problem in a larger class of observers. The latter is obtained by perturbing the Hamiltonian and system-field coupling operators of a linear coherent quantum observer along linear combinations of unitary Weyl operators, whose role here resembles that of the needle variations in the Pontryagin minimum principle. We show that if the observer is a stationary point of the performance functional in the class of linear observers, then it is also a stationary point with respect to the Weyl variations in the larger class of nonlinear observers.

I. INTRODUCTION

The dynamics of a wide class of open quantum systems, interacting with the environment, can be described in the framework of the Hudson-Parthasarathy quantum stochastic calculus [9], [10], [19]. This approach employs quantum stochastic differential equations (QSDEs) which are driven by a quantum mechanical analogue of the classical Wiener process [13]. The quantum Wiener process models a heat reservoir of external fields and acts on a boson Fock space [21], [19]. The drift vector and the dispersion matrix of the QSDEs depend on the system Hamiltonian and the system-field coupling operators. These energy operators specify the evolution of the system as a result of its internal dynamics influenced by the interaction with the environment. The Hamiltonian and the coupling operators are usually modelled as functions of the system variables. In particular, such functions can be polynomials or Weyl quantization integrals [5], [24], which affects the complexity of the resulting quantum system.

An important role in the linear quantum control and filtering theory [3], [12], [22] is played by open quantum harmonic oscillators (OQHOs) [4], [6] with quadratic Hamiltonians and linear system-field coupling operators. The linear-quadratic dependence of the energy operators on the system variables, in combination with the canonical commutation relations (CCRs) between the variables, makes the OQHO dynamics linear (and Gaussian in the case of vacuum fields and Gaussian initial states [11], [20]). Despite some similarities to the classical linear SDEs, the coherent (that is, measurement-free) quantum counterparts [18], [17] to the classical LQG control and filtering problems [1], [14] for OQHOs are complicated by the physical realizability (PR) constraints. The latter are associated with the state-space matrices of the QSDEs for fully quantum controllers or filters and are related, in particular, to the CCR preservation.

We mention one of the existing variational approaches [27], [28] to the coherent quantum LQG (CQLQG) control and coherent quantum filtering (CQF) problems which develops optimality conditions using the Frechet differentiation of the LQG cost with respect to the state-space matrices. The quantum nature of the underlying problem enters this approach only through the PR constraints, with all the other aspects of the method being essentially "classical". The latter has certain advantages, such as practical applicability to the numerical optimization algorithms [25]. However, this approach is limited to linear controllers and filters, and the resulting optimality conditions do not provide insights into whether nonlinear quantum controllers or filters can outperform the linear ones for linear quantum plants.

In the present paper, we consider a CQF problem, similar to [17], [28], where a quantum observer is cascaded in a measurement-free fashion with a linear quantum plant so as to minimize a mean square error with which the observer variables approximate linear combinations of plant variables of interest. Both systems are governed by Markovian Hudson-Parthasarathy QSDEs driven by bosonic fields in vacuum state. We employ a recently proposed fully quantum variational method of [30] based on using a transverse Hamiltonian, an auxiliary time-varying operator which encodes the propagation of perturbations through the unitary system-field evolution. We apply the transverse Hamiltonian approach to the development of first-order necessary conditions of optimality for the CQF problem in a larger class of observers. The latter is obtained by perturbing the Hamiltonian and system-field coupling operators of a linear coherent quantum observer along linear combinations of the unitary Weyl operators [5]. Similar trigonometric polynomials of quantum variables have recently been

†UNSW Canberra, Australia. igor.g.vladimirov@gmail.com.

used in [24] for modelling the uncertainty in system Hamiltonians. In the present paper, the Weyl variations play a different role which resembles that of the needle variations in the proof of the Pontryagin minimum principle [23]. We show that if the observer is a stationary point of the cost functional in the class of linear observers, then it is also a stationary point with respect to the Weyl variations (with the latter leading to nonlinear observers). Therefore, in the mean square optimal CQF problem for linear quantum plants, linear coherent quantum observers are locally sufficient at least in the sense of the Weyl variations of the energy operators.

The paper is organised as follows. Section II outlines notation used in the paper. Section III specifies the class of open quantum stochastic plants being considered. Section IV formulates the mean square optimal CQF problem. Section V employs the transverse Hamiltonian approach in order to study the sensitivity of the performance criterion to perturbations in the energy operators of observers. Section VI applies these results to liner-quadratic perturbations of the observer energy operators and obtains the first-order necessary conditions of optimality in the class of linear observers. Section VII introduces the Weyl variations of the energy operators and establishes the main result of the paper that the stationarity of the quadratic cost functional with respect to linear-quadratic perturbations of a linear observer implies the stationarity with respect to the Weyl variations. Section VIII provides concluding remarks.

II. PRINCIPAL NOTATION

In what follows, $[A,B] := AB - BA$ denotes the commutator of linear operators A and B on a common space. As a linear superoperator, the commutator with a fixed operator A, is denoted by $\text{ad}_A(\cdot) := [A, \cdot]$. This extends to the commutator $(n \times m)$-matrix $[X, Y^{\mathrm{T}}] := XY^{\mathrm{T}} - (YX^{\mathrm{T}})^{\mathrm{T}} = ([X_j, Y_k])_{1 \leqslant j \leqslant n, 1 \leqslant k \leqslant m}$ for a vector X of operators $X_1, \ldots, X_n$ and a vector Y of operators $Y_1, \ldots, Y_m$. Vectors are organized as columns unless indicated otherwise, and the transpose $(\cdot)^{\mathrm{T}}$ acts on matrices of operators as if their entries were scalars. In application to such matrices, $(\cdot)^{\dagger} := ((\cdot)^{\#})^{\mathrm{T}}$ denotes the transpose of the entry-wise operator adjoint $(\cdot)^{\#}$. For complex matrices, $(\cdot)^{\dagger}$ is the usual complex conjugate transpose $(\cdot)^{*} := \overline{((\cdot))}^{\mathrm{T}}$. The subspaces of real symmetric, real antisymmetric and complex Hermitian matrices of order n are denoted by $\mathbb{S}_n$, $\mathbb{A}_n$ and $\mathbb{H}_n := \mathbb{S}_n + i\mathbb{A}_n$, respectively, where $i := \sqrt{-1}$ is the imaginary unit. The real and imaginary parts of a complex matrix are denoted by $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$. These extend to matrices M with operator-valued entries as $\text{Re}M = \frac{1}{2}(M + M^{\#})$ and $\text{Im}M = \frac{1}{2i}(M - M^{\#})$ which consist of self-adjoint operators. Also, $\mathbf{S}(M) := \frac{1}{2}(M + M^{\mathrm{T}})$ denotes the symmetrizer of square matrices. Positive (semi-)definiteness of matrices and the corresponding partial ordering are denoted by $(\succcurlyeq) \succ$. Also, $\mathbb{S}_n^{+}$ and $\mathbb{H}_n^{+}$ denote the sets of positive semi-definite real symmetric and complex Hermitian matrices of order n, respectively. The tensor product of spaces or operators (in particular, the Kronecker product of matrices) is denoted by $\otimes$. The tensor product $A \otimes B$ of operators A and B acting on different spaces will sometimes be abbreviated as AB. The identity matrix of order n is denoted by I_n, while the identity operator on a space H is denoted by $\mathscr{I}_H$. The Frobenius inner product of real or complex matrices is denoted by $\langle M, N \rangle := \text{Tr}(M^{*}N)$. Also, $\|v\|_K := \sqrt{v^{\mathrm{T}}Kv}$ denotes the Euclidean (semi-)norm of a real vector v associated with a real positive (semi-)definite symmetric matrix K. The expectation $\mathbf{E}\xi := \text{Tr}(\rho\xi)$ of a quantum variable ξ over a density operator ρ extends entrywise to vectors and matrices of such variables.

III. QUANTUM PLANTS BEING CONSIDERED

We consider a quantum plant which is modelled as a quantum stochastic system interacting with m external boson fields. The plant has n dynamic variables $X_1(t), \ldots, X_n(t)$ which evolve in time $t \geqslant 0$. The plant variables are self-adjoint operators on a composite plant-field Hilbert space $\mathscr{H} \otimes \mathscr{F}$, where $\mathscr{H}$ is the initial complex separable Hilbert space of the plant which provides a domain for $X_1(0), \ldots, X_n(0)$, and $\mathscr{F}$ is a boson Fock space [19] for the action of quantum Wiener processes $W_1(t), \ldots, W_m(t)$. The latter are self-adjoint operators which model the external boson fields. The energetics of the plant-field interaction is specified by the plant Hamiltonian $H(t)$ and the plant-field coupling operators $L_1(t), \ldots, L_m(t)$ which are self-adjoint operators, representable as time-invariant functions (for example, polynomials with constant coefficients or Weyl quantization integrals [5]) of the plant variables $X_1(t), \ldots, X_n(t)$. Therefore, both $H(0)$ and $L_1(0), \ldots, L_m(0)$ act on the initial space $\mathscr{H}$. Omitting the time arguments, we assemble the plant and field variables, and the coupling operators into vectors:

$$X := \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \qquad W := \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix}, \qquad L := \begin{bmatrix} L_1 \\ \vdots \\ L_m \end{bmatrix}. \tag{1}$$

For what follows, the plant is assumed to be an OQHO [4]. More precisely, the plant variables satisfy the Weyl CCRs $e^{i(u+v)^{\mathrm{T}}X} = e^{iu^{\mathrm{T}}\Theta v}e^{iu^{\mathrm{T}}X}e^{iv^{\mathrm{T}}X}$ for all $u, v \in \mathbb{R}^n$, where $e^{iu^{\mathrm{T}}X}$ is the unitary Weyl operator [5] on $\mathscr{H} \otimes \mathscr{F}$ which is associated with (and inherits time dependence from) the plant variables. The Heisenberg infinitesimal form of the Weyl CCRs is described by

$$[X, X^{\mathrm{T}}] = 2i\Theta \tag{2}$$

where the CCR matrix $\Theta \in \mathbb{A}_n$ represents $\Theta \otimes \mathscr{I}_{\mathscr{H} \otimes \mathscr{F}}$ and remains unchanged. The Hamiltonian of the OQHO is a quadratic polynomial of the plant variables, and the plant-field coupling operators in (1) are linear functions of the variables:

$$H = \frac{1}{2}X^{\mathrm{T}}RX, \qquad L = NX, \tag{3}$$

where $R \in \mathbb{S}_n$ is the energy matrix, and $N \in \mathbb{R}^{m \times n}$ is the coupling matrix. The plant and the external fields form an isolated quantum system whose evolution is described by a unitary operator $U(t)$ on $\mathscr{H} \otimes \mathscr{F}$ driven by the fields and their interaction with the plant according to the QSDE [10], [19]

$$\begin{aligned}
\mathrm{d}U &= -\left(i(H_0\mathrm{d}t + L_0^{\mathrm{T}}\mathrm{d}W) + \frac{1}{2}L_0^{\mathrm{T}}\Omega L_0\mathrm{d}t\right)U \\
&= -U\left(i(H\mathrm{d}t + L^{\mathrm{T}}\mathrm{d}W) + \frac{1}{2}L^{\mathrm{T}}\Omega L\mathrm{d}t\right),
\end{aligned} \tag{4}$$

with initial condition $U_0 := \mathscr{I}_{\mathscr{H} \otimes \mathscr{F}}$. The subscript $(\cdot)_0$ indicates the initial values of time-varying operators (or vectors and matrices thereof), so that $H_0 := H(0)$, $L_0 := L(0)$ and $U_0 := U(0)$, and the time arguments will often be omitted for brevity. Due to the continuous tensor product structure of the Fock space [21], the future-pointing increments $\mathrm{d}W$ commute with adapted processes (including U) taken at the same (or an earlier) moment of time. The matrix $\Omega := (\omega_{jk})_{1 \leqslant j,k \leqslant m} \in \mathbb{H}_m^+$ in (4) is the Ito matrix of the quantum Wiener process W:

$$\mathrm{d}W\mathrm{d}W^{\mathrm{T}} = \Omega\mathrm{d}t, \qquad \Omega := I_m + iJ, \tag{5}$$

and the matrix $J \in \mathbb{A}_m$ specifies the cross-commutations between the entries $W_1, \dots, W_m$ of W:

$$[\mathrm{d}W, \mathrm{d}W^{\mathrm{T}}] = 2iJ\mathrm{d}t, \quad J := \mathbf{J} \otimes I_{m/2}, \quad \mathbf{J} := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{6}$$

where the dimension m is assumed to be even. The QSDE (4) corresponds to an important particular case of open quantum dynamics when the scattering matrix is the identity matrix, and there is no photon exchange between the fields, thus eliminating the gauge processes [19] from consideration. A plant operator σ_0 on the initial space $\mathscr{H}$ (which can be identified with its extension $\sigma \otimes \mathscr{I}_{\mathscr{F}}$ to the plant-field space $\mathscr{H} \otimes \mathscr{F}$) evolves to an operator $\sigma(t)$ on $\mathscr{H} \otimes \mathscr{F}$ at time $t \geqslant 0$ according to the flow

$$\sigma(t) := \mathrm{j}_t(\sigma_0) = U(t)^\dagger(\sigma_0 \otimes \mathscr{I}_{\mathscr{F}})U(t). \tag{7}$$

When it is applied to vectors and matrices of operators, the flow j_t acts entrywise. In view of the identity $(\sigma_0 \otimes \mathscr{I}_{\mathscr{F}})U = U\sigma$, which follows from (7) and the unitarity of U, the second equality in (4) employs the representation of the Hamiltonian and the coupling operators in terms of the flow:

$$H(t) = \mathrm{j}_t(H_0), \qquad L(t) = \mathrm{j}_t(L_0).$$

Note that the flow j_t depends on the energy operators H_0 and L_0 (or the energy and coupling matrices R and N in (3) in the case of OQHOs). More precisely, any perturbation of H_0 and L_0 (as functions of the fixed set of system variables X_0) modifies the flow. Now, the quantum adapted process σ in (7) satisfies the following Hudson-Parthasarathy QSDE [10], [19]:

$$\mathrm{d}\sigma = \mathscr{G}(\sigma)\mathrm{d}t - i[\sigma, L^{\mathrm{T}}]\mathrm{d}W, \quad \mathscr{G}(\sigma) := i[H, \sigma] + \mathscr{D}(\sigma). \tag{8}$$

Here, $\mathscr{D}$ is the Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) decoherence superoperator [7], [15] which acts on σ as

$$\begin{aligned}
\mathscr{D}(\sigma) &:= \frac{1}{2}\left(L^{\mathrm{T}}\Omega[\sigma, L] + [L^{\mathrm{T}}, \sigma]\Omega L\right) \\
&= -[\sigma, L^{\mathrm{T}}]\Omega L - \frac{1}{2}[L^{\mathrm{T}}\Omega L, \sigma] \\
&= -[\sigma, X^{\mathrm{T}}]N^{\mathrm{T}}\Omega NX - \frac{1}{2}[X^{\mathrm{T}}N^{\mathrm{T}}\Omega NX, \sigma].
\end{aligned} \tag{9}$$

The last two equalities in (9) are convenient for the entrywise evaluation of $\mathscr{D}$ at vectors of operators. The superoperator $\mathscr{G}$ in (8) is referred to as the GKSL generator. In application to the vector X of plant variables, the flow (7) acts entrywise as

$$X(t) := \mathrm{j}_t(X_0) = U(t)^\dagger(X_0 \otimes \mathscr{I}_{\mathscr{F}})U(t).$$

In view of (2) and (3), the corresponding QSDE (8) takes the form

$$\mathrm{d}X = \mathscr{G}(X)\mathrm{d}t - i[X, L^{\mathrm{T}}]\mathrm{d}W = AX\mathrm{d}t + B\mathrm{d}W, \tag{10}$$

with the n-dimensional drift vector $\mathscr{G}(X) = AX$ and the dispersion $(n \times m)$-matrix $-i[X, L^{\mathrm{T}}] = -i[X, X^{\mathrm{T}}]N^{\mathrm{T}} = B$, where the matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are given by

$$A := 2\Theta(R + N^{\mathrm{T}}JN), \qquad B := 2\Theta N^{\mathrm{T}}. \tag{11}$$

The interaction of the input field with the plant produces an m-dimensional output field

$$Y(t) := \begin{bmatrix} Y_1(t) \\ \vdots \\ Y_m(t) \end{bmatrix} = U(t)^\dagger (\mathscr{I}_{\mathscr{H}} \otimes W(t))U(t), \tag{12}$$

where the plant-field unitary evolution is applied to the current input field variables (which is closely related to the innovation role of the quantum Wiener process W in the QSDEs). The output field satisfies the QSDE

$$dY = 2JL dt + dW = CX dt + dW, \tag{13}$$

where J is the matrix from (6), L is the vector of plant-field coupling operators in (1) and (3), and the matrix $C \in \mathbb{R}^{m \times n}$ is given by

$$C := 2JN. \tag{14}$$

IV. COHERENT QUANTUM FILTERING PROBLEM

Consider a measurement-free cascade connection of the quantum plant, described in Section III, with another open quantum system. The latter plays the role of a coherent quantum observer and is driven by ΠY (which is part of the plant output Y in (12)) and a quantum Wiener process ω of even dimension μ on a boson Fock space $\mathfrak{F}$; see Fig. 1. The matrix $\Pi \in \mathbb{R}^{p \times m}$
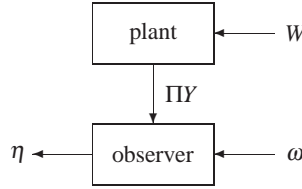

Fig. 1. The cascade connection of a quantum observer with a quantum plant, mediated by the field ΠY and affected by the environment through the quantum Wiener processes W and ω. Also shown is the observer output η.

is formed from conjugate pairs of rows of a permutation matrix of order m, with $p \leqslant m$. The observer has its own initial Hilbert space $\mathfrak{H}$, dynamic variables $\xi_1, \ldots, \xi_v$ with a CCR matrix $\vartheta \in \mathbb{A}_v$, and a $(p + \mu)$-dimensional output field η:

$$\xi := \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_v \end{bmatrix}, \qquad \omega := \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_\mu \end{bmatrix}, \qquad \eta := \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_{p+\mu} \end{bmatrix}. \tag{15}$$

We denote the observer Hamiltonian by Γ, while the vectors of operators of coupling of the observer with the selected plant output ΠY and the quantum Wiener process ω are denoted by

$$\Phi := \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_p \end{bmatrix}, \qquad \Psi := \begin{bmatrix} \Psi_1 \\ \vdots \\ \Psi_\mu \end{bmatrix}, \tag{16}$$

respectively. The Hamiltonian Γ and the coupling operators $\Phi_1, \ldots, \Phi_p$ and $\Psi_1, \ldots, \Psi_\mu$ are functions of the dynamic variables $\xi_1, \ldots, \xi_v$ of the observer and hence, commute with functions of the plant variables, including the plant Hamiltonian H and the plant-field coupling operators in L. The plant and the observer form a composite open quantum stochastic system, whose vector $\mathscr{X}$ of dynamic variables satisfies the CCRs

$$[\mathscr{X}, \mathscr{X}^{\mathrm{T}}] = 2i\Theta, \qquad \Theta := \begin{bmatrix} \Theta & 0 \\ 0 & \vartheta \end{bmatrix}, \qquad \mathscr{X} := \begin{bmatrix} X \\ \xi \end{bmatrix} \tag{17}$$

and is driven by a combined quantum Wiener process $\mathscr{W}$ with the Ito table

$$d\mathscr{W} d\mathscr{W}^{\mathrm{T}} = \Omega dt, \qquad \Omega := \begin{bmatrix} \Omega & 0 \\ 0 & \mho \end{bmatrix}, \qquad \mathscr{W} := \begin{bmatrix} W \\ \omega \end{bmatrix}. \tag{18}$$

Here, $\mho$ is the Ito matrix of the quantum Wiener process ω of the observer which is defined similarly to Ω in (5) and (6):

$$d\omega d\omega^{\mathrm{T}} = \mho dt, \qquad \mho := I_\mu + i\mathbf{J} \otimes I_{\mu/2}. \tag{19}$$

The Hamiltonian $\mathbf{H}$ of the plant-observer system and the vector $\mathbf{L}$ of operators of coupling with $\mathscr{W}$ can be computed by using the quantum feedback network formalism [8] as

$$\mathbf{H} = H + \Gamma + \Phi^{\mathrm{T}} \Pi JL, \qquad \mathbf{L} = \begin{bmatrix} L + \Pi^{\mathrm{T}} \Phi \\ \Psi \end{bmatrix}. \tag{20}$$

4

While the plant dynamics, governed by (10) and (13), remains unaffected by the observer, the dynamic variables of the latter in (15) are governed by the QSDE

$$d\xi = \mathbf{G}(\xi)dt - i[\xi, \mathbf{L}^T]d\mathcal{W}$$
$$= (i[\Gamma, \xi] + \Delta(\xi))dt - i[\xi, \Phi^T]\Pi dY - i[\xi, \Psi^T]d\omega. \tag{21}$$

Here,

$$\mathbf{G}(\zeta) := i[\mathbf{H}, \zeta] + \mathbf{D}(\zeta) \tag{22}$$

denotes the GKSL generator for the plant-observer system, and

$$\mathbf{D}(\zeta) = -[\zeta, \mathbf{L}^T]\mathbf{\Omega L} - \frac{1}{2}[\mathbf{L}^T\mathbf{\Omega L}, \zeta] \tag{23}$$

is the corresponding decoherence superoperator, similar to (9). In (21), use is also made of the partial GKSL decoherence superoperator Δ which acts on the observer variables as

$$\Delta(\xi) = -[\xi, \Phi^T]\Pi\Omega\Pi^T\Phi - [\xi, \Psi^T]\mho\Psi$$
$$- \frac{1}{2}[\Phi^T\Pi\Omega\Pi^T\Phi + \Psi^T\mho\Psi, \xi] \tag{24}$$

in view of (17)–(20), with $\Pi\Omega\Pi^T \in \mathbb{H}_p^+$ being the quantum Ito matrix of ΠW. Now, similarly to [17], [28], we formulate a CQF problem as the minimization of the steady-state mean square discrepancy

$$\mathscr{Z} := \lim_{t \to +\infty} \mathbf{E}Z(t) \longrightarrow \min, \qquad Z := E^T E, \quad E := FX - G\xi \tag{25}$$

between q linear combinations of the plant variables of interest and observer variables as specified by given matrices $F \in \mathbb{R}^{q \times n}$ and $G \in \mathbb{R}^{q \times \nu}$ (with E being interpreted as the estimation error). Here, the quantum expectation $\mathbf{E}(\cdot)$ is taken over the tensor product $\rho := \varpi \otimes \upsilon$ of the initial quantum state ϖ of the plant-observer system and the vacuum state υ in the composite boson Fock space $\mathscr{F} \otimes \mathfrak{F}$ for the external fields. Also, it is assumed that the plant-observer system is ergodic, whereby the limit in (25) reduces to averaging over the invariant state of the system, provided the plant and observer variables satisfy an appropriate version of the uniform mean square integrability condition. The criterion process Z in (25) is a quadratic function of the vector $\mathscr{X}$ from (17):

$$Z := \mathscr{X}^T \mathscr{C}^T \mathscr{C} \mathscr{X}, \qquad \mathscr{C} := \begin{bmatrix} F & -G \end{bmatrix}. \tag{26}$$

The minimization in (25) is carried out over the observer Hamiltonian Γ and the vector Φ of the observer-plant coupling operators in (16), while Ψ and all the dimensions are fixed. This problem extends [17], [28] in that we do not restrict attention to linear observers even though the plant is an OQHO.

V. INFINITESIMAL PERTURBATION ANALYSIS

In order to develop first-order necessary conditions of optimality for the CQF problem (25) in an extended class of observers, we will apply the transverse Hamiltonian variational method of [30] to the infinitesimal perturbation analysis of the performance criterion $\mathscr{Z}$. To this end, suppose Γ_0 and Φ_0 depend smoothly (for example, linearly) on a small scalar parameter ε and are perturbed in the directions

$$K_0 := \Gamma_0', \qquad M_0 := \Phi_0', \tag{27}$$

consisting of self-adjoint operators on the observer initial space $\mathfrak{H}$, representable as functions of the observer variables, with $(\cdot)' := \partial_\varepsilon(\cdot)|_{\varepsilon=0}$. The corresponding perturbations of the plant-observer Hamiltonian and the coupling operators in (20) are

$$\mathbf{H}_0' = K_0 + M_0^T\Pi J L_0, \qquad \mathbf{L}_0' = \begin{bmatrix} \Pi^T M_0 \\ 0 \end{bmatrix}. \tag{28}$$

The propagation of these perturbations of the energy operators through the unitary evolution $\mathbf{U}$ of the plant-observer-field system on the space $\mathscr{H} \otimes \mathfrak{H} \otimes \mathscr{F} \otimes \mathfrak{F}$ is encoded by the transverse Hamiltonian [30]. The latter is a time-varying self-adjoint operator defined by $Q := i\mathbf{U}^\dagger \mathbf{U}'$, so that $\mathbf{U}(t)' = -i\mathbf{U}(t)Q(t)$ for all $t \geq 0$, with zero initial condition $Q_0 = 0$. In view of (18), (27) and (28), the general QSDE obtained in [30, Theorem 1] for the transverse Hamiltonian, takes the form:

$$dQ = \left(K + M^T\Pi J L - \mathrm{Im}\left(\mathbf{L}^T\mathbf{\Omega}\begin{bmatrix} \Pi^T M \\ 0 \end{bmatrix}\right)\right)dt + \begin{bmatrix} M^T\Pi & 0 \end{bmatrix}d\mathcal{W}$$
$$= \left(K - \mathrm{Im}((2L + \Pi^T\Phi)^T\Omega\Pi^T M)\right)dt + M^T\Pi dW. \tag{29}$$

Here, $K := \mathbf{j}_t(K_0)$ and $M := \mathbf{j}_t(M_0)$ are the evolved versions of the initial perturbations from (27) under the unperturbed flow $\mathbf{j}_t(\zeta) := \mathbf{U}(t)^\dagger(\mathscr{I}_\mathscr{H} \otimes \zeta \otimes \mathscr{I}_{\mathscr{F} \otimes \mathfrak{F}})\mathbf{U}(t)$ of the plant-observer-field system (which is applied here to observer operators ζ on $\mathfrak{H}$). In (29), use is also made of the relation $\mathrm{Im}(L^T\Omega\Pi^T M) = -M^T\Pi J L$ which follows from (5), the commutativity

$[L, M^{\mathrm{T}}] = 0$ and the antisymmetry of J. Since the criterion process Z in (25) does not depend explicitly on the energy operators, the corresponding formal Gateaux derivative $\mathscr{Z}'$ of the cost functional can be computed by using [30, Theorem 2, Section VII] as

$$\mathscr{Z}' := \lim_{t \to +\infty} \mathbf{E}\phi(t), \qquad \phi := i[Q, Z]. \tag{30}$$

Here, ϕ is the derivative process [30] associated with Z. Its expectation satisfies the integro-differential equation

$$(\mathbf{E}\phi)^{\cdot} = i\mathbf{E}[Q, \mathbf{G}(Z)] + \mathbf{E}\chi(Z), \tag{31}$$

where $\mathbf{G}$ is the unperturbed plant-observer GKSL generator given by (20), (22), (23), and χ is an auxiliary linear superoperator acting on plant-observer system operators σ as

$$\chi(\sigma) := i[K - \mathrm{Im}((2L + \Pi^{\mathrm{T}}\Phi)^{\mathrm{T}}\Omega\Pi^{\mathrm{T}}M), \sigma]$$
$$- 2\mathrm{Re}([\sigma, (L + \Pi^{\mathrm{T}}\Phi)^{\mathrm{T}}]\Omega\Pi^{\mathrm{T}}M). \tag{32}$$

Note that $\chi(\sigma)$ depends linearly on the perturbations K and M. Now, in addition to the plant being an OQHO, suppose the unperturbed observer is also an OQHO with energy matrix $r \in \mathbb{S}_\nu$ and coupling matrices $N_1 \in \mathbb{R}^{p \times \nu}$ and $N_2 \in \mathbb{R}^{\mu \times \nu}$. The corresponding observer Hamiltonian Γ and the coupling operators in (16) are

$$\Gamma = \frac{1}{2}\xi^{\mathrm{T}}r\xi, \qquad \Phi = N_1\xi, \qquad \Psi = N_2\xi. \tag{33}$$

In this case, in view of (18) and (24), the QSDE (21) becomes linear:

$$\mathrm{d}\xi = a\xi\mathrm{d}t + b_1\mathrm{d}Y + b_2\mathrm{d}\omega, \tag{34}$$

where the matrices $a \in \mathbb{R}^{\nu \times \nu}$, $b_1 \in \mathbb{R}^{\nu \times p}$ and $b_2 \in \mathbb{R}^{\nu \times \mu}$ are computed as

$$a := 2\vartheta(r + N_1^{\mathrm{T}}\Pi J\Pi^{\mathrm{T}}N_1 + N_2^{\mathrm{T}}\mathrm{Im}\mho N_2), \tag{35}$$
$$b_1 := 2\vartheta N_1^{\mathrm{T}}\Pi, \qquad b_2 := 2\vartheta N_2^{\mathrm{T}}. \tag{36}$$

Therefore, the plant-observer system is governed by a linear QSDE

$$\mathrm{d}\mathscr{X} = \mathscr{A}\mathscr{X}\mathrm{d}t + \mathscr{B}\mathrm{d}\mathscr{W}, \quad \mathscr{A} := \begin{bmatrix} A & 0 \\ b_1C & a \end{bmatrix}, \quad \mathscr{B} := \begin{bmatrix} B & 0 \\ b_1 & b_2 \end{bmatrix}. \tag{37}$$

For what follows, both matrices A and a are assumed to be Hurwitz, and hence, so is $\mathscr{A}$. This implies that the plant-observer system is ergodic and has a unique invariant state which is Gaussian [20] with zero mean and quantum covariance matrix $\mathscr{P} + i\Theta \in \mathbb{H}_{n+\nu}^+$. Here, $\mathscr{P} \in \mathbb{S}_{n+\nu}^+$ is the controllability Gramian [1], [14] of the pair $(\mathscr{A}, \mathscr{B})$ which is a unique solution of the ALE

$$\mathscr{A}\mathscr{P} + \mathscr{P}\mathscr{A}^{\mathrm{T}} + \mathscr{B}\mathscr{B}^{\mathrm{T}} = 0. \tag{38}$$

Lemma 1: Suppose the plant and the unperturbed observer are OQHOs described by (3), (10)–(14) and (33)–(36), with Hurwitz matrices A and a. Also, let the observer Hamiltonian and the observer-plant coupling operators are perturbed according to (27). Then the corresponding formal Gateaux derivative of the cost functional in (25) can be computed as

$$\mathscr{Z}' = \mathbf{E}_*\chi(\sigma), \qquad \sigma := \mathscr{X}^{\mathrm{T}}\mathscr{Q}\mathscr{X}. \tag{39}$$

Here, $\mathscr{Q} \in \mathbb{S}_{n+\nu}^+$ is the observability Gramian of the pair $(\mathscr{A}, \mathscr{C})$ satisfying the ALE

$$\mathscr{A}^{\mathrm{T}}\mathscr{Q} + \mathscr{Q}\mathscr{A} + \mathscr{C}^{\mathrm{T}}\mathscr{C} = 0, \tag{40}$$

where the matrices $\mathscr{C}$ and $\mathscr{A}$ are given by (26) and (37). Also, $\mathbf{E}_*(\cdot)$ denotes the quantum expectation over the invariant Gaussian state of the plant-observer system (with $\mathbf{E}_*\mathscr{X} = 0$ and $\mathscr{P} := \mathrm{Re}\mathbf{E}_*(\mathscr{X}\mathscr{X}^{\mathrm{T}})$ found from (38)). Furthermore, the superoperator χ from (32) acts on the operator σ in (39) as

$$\chi(\sigma) = i\left[K - \mathrm{Im}\left(\mathscr{X}^{\mathrm{T}}\begin{bmatrix} 2N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}M\right), \sigma\right]$$
$$+ 8\mathrm{Im}\left(\mathscr{X}^{\mathrm{T}}\mathscr{Q}\Theta\begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}M\right). \tag{41}$$

Proof: Since Z in (26) is a quadratic form of $\mathscr{X}$, then (30) and (31) imply that

$$\mathscr{Z}' = \langle \mathscr{C}^{\mathrm{T}}\mathscr{C}, \Upsilon \rangle, \qquad \Upsilon := i\lim_{t \to +\infty}\mathbf{E}[Q, \Xi], \tag{42}$$

6

where $\Upsilon \in \mathbb{S}_{n+v}$ is the limit mean value of the derivative process $i[Q, \Xi]$ associated with $\Xi := \mathrm{Re}(\mathscr{X}\mathscr{X}^{\mathrm{T}}) = \mathscr{X}\mathscr{X}^{\mathrm{T}} - i\Theta$. The GKSL generator $\mathbf{G}$ of the unperturbed linear plant-observer system acts on Ξ as $\mathbf{G}(\Xi) = \mathscr{A}\Xi + \Xi\mathscr{A}^{\mathrm{T}} + \mathscr{B}\mathscr{B}^{\mathrm{T}}$, and hence, similarly to [30, Example 3, Section VII], the matrix Υ satisfies the ALE

$$\mathscr{A}\Upsilon + \Upsilon\mathscr{A}^{\mathrm{T}} + \mathbf{E}_* \chi(\Xi) = 0, \tag{43}$$

where the superoperator χ in (32) is applied to Ξ entrywise. By combining the integral representations of the solutions of the ALEs (40), (43) and using duality, it follows from (42) that

$$
\begin{aligned}
\mathscr{L}' &= \left\langle \mathscr{C}^{\mathrm{T}}\mathscr{C}, \int_0^{+\infty} \mathrm{e}^{t\mathscr{A}} \mathbf{E}_* \chi(\Xi) \mathrm{e}^{t\mathscr{A}^{\mathrm{T}}} \mathrm{d}t \right\rangle \\
&= \left\langle \int_0^{+\infty} \mathrm{e}^{t\mathscr{A}^{\mathrm{T}}} \mathscr{C}^{\mathrm{T}}\mathscr{C} \mathrm{e}^{t\mathscr{A}} \mathrm{d}t, \mathbf{E}_* \chi(\Xi) \right\rangle \\
&= \langle \mathscr{Q}, \mathbf{E}_* \chi(\Xi) \rangle = \mathbf{E}_* \chi(\mathscr{X}^{\mathrm{T}} \mathscr{Q} \mathscr{X}),
\end{aligned}
$$

which establishes (39) since $\chi(\Xi) = \chi(\mathscr{X}\mathscr{X}^{\mathrm{T}})$. Indeed, σ enters $\chi(\sigma)$ in (32) only through the commutators $[\sigma, \cdot]$ and hence, $\chi(\sigma + \theta) = \chi(\sigma)$ for any $\theta \in \mathbb{C}$. The representation (41) is obtained by substituting L from (3), Φ, Ψ from (33) and σ from (39) into (32), and using the relation $\frac{i}{2}[\sigma, \mathscr{X}] = 2\Theta\mathscr{Q}\mathscr{X}$. In particular, the second line of (41) is established by

$$
\begin{aligned}
-2\mathrm{Re}([\sigma, (L + \Pi^{\mathrm{T}}\Phi)^{\mathrm{T}}]\Omega\Pi^{\mathrm{T}}M) & \\
&= -2\mathrm{Re}\left([\sigma, \mathscr{X}^{\mathrm{T}}]\begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}M\right) \\
&= -2\mathrm{Re}\left((-4i\Theta\mathscr{Q}\mathscr{X})^{\mathrm{T}}\begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}M\right) \\
&= 8\mathrm{Im}\left(\mathscr{X}^{\mathrm{T}}\mathscr{Q}\Theta\begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}M\right).
\end{aligned}
$$

■

Although the unperturbed observer in Lemma 1 is an OQHO, the perturbations (27) are not assumed to be linear-quadratic. Therefore, the lemma provides a perturbative tool to develop conditions for such an observer to be a stationary point of the CQF problem (25) in a wider class of observers.

VI. OPTIMALITY AMONG LINEAR OBSERVERS

We will now apply Lemma 1 to the first-order necessary conditions of optimality among linear observers. To this end, associated with the Gramians $\mathscr{P}$ and $\mathscr{Q}$ from (38) and (40) is the Hankelian

$$\mathscr{E} := \begin{bmatrix} \mathscr{E}_{11} & \mathscr{E}_{12} \\ \mathscr{E}_{21} & \mathscr{E}_{22} \end{bmatrix} := \mathscr{Q}\mathscr{P} \tag{44}$$

which, together with $\mathscr{P}$ and $\mathscr{Q}$, is split into appropriately dimensioned blocks $(\cdot)_{jk}$ according to the partitioning of $\mathscr{X}$ in (17), with $(\cdot)_{j\bullet}$ the jth block-row and $(\cdot)_{\bullet k}$ the kth block-column.

Theorem 1: Under the assumptions of Lemma 1, the linear observer, described in the lemma, is a stationary point of the CQF problem (25) in the class of linear observers if and only if

$$\vartheta \mathscr{E}_{22} \in \mathbb{A}_v, \tag{45}$$

$$\Pi(C\mathscr{E}_{21}^{\mathrm{T}} + B^{\mathrm{T}}\mathscr{Q}_{12} + b_1^{\mathrm{T}}\mathscr{Q}_{22})\vartheta = \Pi J b_1^{\mathrm{T}}\mathscr{E}_{22}. \tag{46}$$

Proof: Let the matrices r and N_1 in (33) be smooth functions of the parameter ε, and hence, the corresponding perturbations of the linear observer in (27) take the form

$$K = \frac{1}{2}\xi^{\mathrm{T}}r'\xi, \qquad M = N_1'\xi, \tag{47}$$

where the matrices $r' \in \mathbb{S}_v$ and $N_1' \in \mathbb{R}^{p \times v}$ can be arbitrary. Substitution of (47) into (41) leads to

$$
\begin{aligned}
\chi(\sigma) =& i\left[\frac{1}{2}\xi^{\mathrm{T}}r'\xi - \mathrm{Im}\left(\mathscr{X}^{\mathrm{T}}\begin{bmatrix} 2N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}N_1'\xi\right), \sigma\right] \\
&+ 8\mathrm{Im}\left(\mathscr{X}^{\mathrm{T}}\mathscr{Q}\Theta\begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}N_1'\xi\right),
\end{aligned} \tag{48}
$$

where σ is the quantum variable from (39). The identity $\frac{i}{2}[\mathscr{X}^{\mathrm{T}}R_1\mathscr{X}, \mathscr{X}^{\mathrm{T}}R_2\mathscr{X}] = 2\mathscr{X}^{\mathrm{T}}(R_2\Theta R_1 - R_1\Theta R_2)\mathscr{X}$, which holds for any $R_1, R_2 \in \mathbb{S}_{n+v}$ due to the CCRs (17), implies that

$$\frac{i}{2}[\xi^{\mathrm{T}}r'\xi, \sigma] = 2\mathscr{X}^{\mathrm{T}}\left(\mathscr{Q}\Theta\begin{bmatrix} 0 & 0 \\ 0 & r' \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & r' \end{bmatrix}\Theta\mathscr{Q}\right)\mathscr{X}$$
$$= 2\mathscr{X}^{\mathrm{T}}\left(\begin{bmatrix} 0 & \mathscr{Q}_{\bullet 2}\vartheta r' \end{bmatrix} - \begin{bmatrix} 0 \\ r'\vartheta\mathscr{Q}_{2\bullet} \end{bmatrix}\right)\mathscr{X}.$$

The averaging of the latter representation over the invariant quantum state leads to

$$\frac{i}{2}\mathbf{E}_*[\xi^{\mathrm{T}}r'\xi, \sigma] = 2\left\langle \begin{bmatrix} 0 & \mathscr{Q}_{\bullet 2}\vartheta r' \end{bmatrix} - \begin{bmatrix} 0 \\ r'\vartheta\mathscr{Q}_{2\bullet} \end{bmatrix}, \mathscr{P} \right\rangle$$
$$= -4\left\langle \vartheta\mathscr{Q}_{2\bullet}\mathscr{P}_{\bullet 2}, r' \right\rangle = -4\left\langle \mathbf{S}(\vartheta\mathscr{E}_{22}), r' \right\rangle,$$

where the equality $\mathscr{Q}_{2\bullet}\mathscr{P}_{\bullet 2} = \mathscr{E}_{22}$ follows from (44), and the symmetry of r' is used. Therefore, in view of (39), the corresponding formal Frechet derivative $\partial_r\mathscr{Z} = -4\mathbf{S}(\vartheta\mathscr{E}_{22})$ of the cost functional vanishes if and only if the matrix $\mathscr{E}_{22}$ satisfies (45). By a similar reasoning, in view of (48), the Gateaux derivative of $\mathscr{Z}$ along M in (47) takes the form

$$8\mathrm{Im}\mathbf{E}_*\left(\mathscr{X}^{\mathrm{T}}\mathscr{Q}\Theta\begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}N_1'\xi\right)$$

$$- \mathrm{Re}\mathbf{E}_*\left[\mathscr{X}^{\mathrm{T}}\begin{bmatrix} 2N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\Omega\Pi^{\mathrm{T}}N_1'\xi, \sigma\right]$$

$$= 8\mathrm{Im}\left\langle \mathscr{Q}\Theta\begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\overline{\Omega}\Pi^{\mathrm{T}}N_1', \mathscr{P}_{\bullet 2} + \begin{bmatrix} 0 \\ i\vartheta \end{bmatrix} \right\rangle$$

$$+ 4\mathrm{Re}\left\langle \begin{bmatrix} 0 & \begin{bmatrix} 2N^{\mathrm{T}} \\ N_1^{\mathrm{T}}\Pi \end{bmatrix}\overline{\Omega}\Pi^{\mathrm{T}}N_1' \end{bmatrix}, i(\mathscr{E}^{\mathrm{T}}\Theta - \Theta\mathscr{E}) \right\rangle$$

$$= 8\left\langle \Pi J\begin{bmatrix} N & \Pi^{\mathrm{T}}N_1 \end{bmatrix}\Theta\mathscr{E}_{\bullet 2} - \Pi\begin{bmatrix} N & \Pi^{\mathrm{T}}N_1 \end{bmatrix}\Theta\mathscr{Q}_{\bullet 2}\vartheta \right.$$

$$\left. + \Pi JN(\mathscr{E}_{21}^{\mathrm{T}}\vartheta - \Theta\mathscr{E}_{12}) - \Pi J\Pi^{\mathrm{T}}N_1\mathbf{S}(\vartheta\mathscr{E}_{22}), N_1' \right\rangle$$

$$= 4\left\langle \Pi((C\mathscr{E}_{21}^{\mathrm{T}} + B^{\mathrm{T}}\mathscr{Q}_{12} + b_1^{\mathrm{T}}\mathscr{Q}_{22})\vartheta - Jb_1^{\mathrm{T}}\mathscr{E}_{22}), N_1' \right\rangle, \tag{49}$$

where use is also made of (11), (14), (36) and (44). Here, $\mathbf{S}(\vartheta\mathscr{E}_{22}) = 0$ under the condition (45), in which case, the formal Frechet derivative $\partial_{N_1}\mathscr{Z} = 4\Pi((C\mathscr{E}_{21}^{\mathrm{T}} + B^{\mathrm{T}}\mathscr{Q}_{12} + b_1^{\mathrm{T}}\mathscr{Q}_{22})\vartheta - Jb_1^{\mathrm{T}}\mathscr{E}_{22})$ vanishes if and only if (46) holds. Therefore, (45) and (46) are indeed equivalent to the stationarity in the class of linear observers. ∎

VII. WEYL VARIATIONS OF OBSERVERS

Now, consider the following bounded perturbations of the observer Hamiltonian and the observer-plant coupling operators in (27):

$$K := \mathrm{Re}(\alpha\mathsf{W}_u) = \frac{1}{2}(\alpha\mathsf{W}_u + \overline{\alpha}\mathsf{W}_{-u}) = |\alpha|\cos(u^{\mathrm{T}}\xi + \gamma), \tag{50}$$

$$M := \mathrm{Re}(\beta\mathsf{W}_u) = \frac{1}{2}(\beta\mathsf{W}_u + \overline{\beta}\mathsf{W}_{-u}) = (|\beta_k|\cos(u^{\mathrm{T}}X + \theta_k))_{1\leqslant k\leqslant p}, \tag{51}$$

with parameters $\alpha \in \mathbb{C}$, $\beta := (\beta_k)_{1\leqslant k\leqslant p} \in \mathbb{C}^p$, $\gamma := \arg\alpha$, $\theta_k := \arg\beta_k$ and $u \in \mathbb{R}^v$. Here, $\mathsf{W}_u := e^{iu^{\mathrm{T}}\xi} = \mathsf{W}_{-u}^\dagger$ is the unitary Weyl operator associated with the observer variables. We will refer to the perturbations (50) and (51) as the Weyl variations (of the observer energy operators). Similar trigonometric polynomials of quantum variables have recently been used in [24] to model uncertainties in system Hamiltonians. However, in what follows, the Weyl variations play a different role which resembles that of the needle variations in the proof of the Pontryagin minimum principle [23].

Theorem 2: Under the assumptions of Lemma 1, the linear observer satisfies the first-order necessary conditions of optimality in the CQF problem (25) among linear observers if and only if it is a stationary point with respect to the Weyl variations in (50) and (51) for any $\alpha \in \mathbb{C}$, $\beta \in \mathbb{C}^p$, $u \in \mathbb{R}^v$.

Proof: The stationarity of the cost functional $\mathscr{Z}$ with respect to arbitrary Weyl variations of the observer implies the stationarity with respect to the linear-quadratic perturbations (47) because $\partial_u\mathsf{W}_u\big|_{u=0} = i\xi$ and $\partial_u^2\mathsf{W}_u\big|_{u=0} = -\mathrm{Re}(\xi\xi^{\mathrm{T}})$. In fact, arbitrary polynomial perturbations can be reproduced by the Weyl variations which form a larger class of perturbations. Therefore, it remains to prove that (45) and (46) imply the stationarity with respect to arbitrary Weyl variations. By Lemma 1, the linear response of the cost functional $\mathscr{Z}$ to the Weyl variation K of the observer Hamiltonian in (50) takes the form

$$i\mathbf{E}_*[\mathrm{Re}(\alpha\mathsf{W}_u), \sigma] = -\mathrm{Im}(\alpha\mathbf{E}_*[\mathsf{W}_u, \sigma]). \tag{52}$$

8

Let S_u denote a linear superoperator, which is parameterized by $u \in \mathbb{R}^\nu$ and carries out a unitary similarity transformation of quantum variables as $\mathsf{S}_u(\zeta) := \mathsf{W}_u \zeta \mathsf{W}_{-u} = \mathrm{e}^{i \mathrm{ad}_{u^{\mathrm{T}} \xi}}(\zeta) = \sum_{k=0}^{+\infty} \frac{i^k}{k!} \mathrm{ad}_{u^{\mathrm{T}} \xi}^k(\zeta)$, where use is made of a well-known identity for operator exponentials [16], [31]. Its application to $\mathscr{X}$ from (17) leads to $\mathsf{S}_u(\mathscr{X}) = \mathscr{X} + \begin{bmatrix} 0 \\ 2\vartheta u \end{bmatrix}$ since $[X, \xi^{\mathrm{T}}] = 0$ (whence $\mathsf{S}_u(X) = X$) and $[u^{\mathrm{T}}\xi, \xi] = -[\xi, u^{\mathrm{T}}\xi] = -[\xi, \xi^{\mathrm{T}}]u = -2i\vartheta u$ (whereby $\mathsf{S}_u(\xi) = \xi + 2\vartheta u$). Therefore, in view of (39), $\mathsf{S}_u(\sigma) = \mathsf{S}_u(\mathscr{X})^{\mathrm{T}} \mathscr{Q} \mathsf{S}_u(\mathscr{X}) = \sigma - 4u^{\mathrm{T}} \vartheta (\mathscr{Q}_{2\bullet} \mathscr{X} + \mathscr{Q}_{22} \vartheta u)$, and hence,

$$
\begin{aligned}
[\mathsf{W}_u, \sigma] &= (\mathsf{S}_u(\sigma) - \sigma) \mathsf{W}_u \\
&= -4u^{\mathrm{T}} \vartheta (\mathscr{Q}_{21} X + \mathscr{Q}_{22}(\xi + \vartheta u)) \mathsf{W}_u \\
&= 4i u^{\mathrm{T}} \vartheta (\mathscr{Q}_{21} \partial_w + \mathscr{Q}_{22} \partial_u) \mathrm{e}^{i(w^{\mathrm{T}} X + u^{\mathrm{T}} \xi)} \big|_{w=0}.
\end{aligned}
\tag{53}
$$

Here, $\mathrm{e}^{i w^{\mathrm{T}} X} \mathsf{W}_u = \mathrm{e}^{i(w^{\mathrm{T}} X + u^{\mathrm{T}} \xi)}$ in view of the commutativity $[\mathrm{e}^{i w^{\mathrm{T}} X}, \mathsf{W}_u] = 0$ between the Weyl operators for all $w \in \mathbb{R}^n$, $u \in \mathbb{R}^\nu$, and hence,

$$
X \mathsf{W}_u = -i \partial_w \mathrm{e}^{i w^{\mathrm{T}} X} \big|_{w=0} \mathsf{W}_u = -i \partial_w \mathrm{e}^{i(w^{\mathrm{T}} X + u^{\mathrm{T}} \xi)} \big|_{w=0}.
\tag{54}
$$

Also, the relation $(\xi + \vartheta u) \mathsf{W}_u = -i \partial_u \mathsf{W}_u$ follows from

$$
\begin{aligned}
\xi \mathsf{W}_u &= -i \partial_v \mathsf{W}_v \big|_{v=0} \mathsf{W}_u \\
&= -i \partial_v (\mathrm{e}^{i u^{\mathrm{T}} \vartheta v} \mathsf{W}_{u+v}) \big|_{v=0} = -\vartheta u \mathsf{W}_u - i \partial_u \mathsf{W}_u
\end{aligned}
\tag{55}
$$

due to the Weyl CCRs $\mathsf{W}_v \mathsf{W}_u = \mathrm{e}^{i u^{\mathrm{T}} \vartheta v} \mathsf{W}_{u+v}$ which hold for all $u, v \in \mathbb{R}^\nu$. By averaging (53) over the invariant Gaussian quantum state, whose quasi-characteristic function [2] is $\mathbf{E}_* \mathrm{e}^{i \lambda^{\mathrm{T}} \mathscr{X}} = \mathrm{e}^{-\frac{1}{2} \|\lambda\|_{\mathscr{P}}^2}$ for any $\lambda := \begin{bmatrix} w \\ u \end{bmatrix} \in \mathbb{R}^{n+\nu}$, it follows that

$$
\begin{aligned}
\mathbf{E}_* [\mathsf{W}_u, \sigma] &= 4i u^{\mathrm{T}} \vartheta \mathscr{Q}_{2\bullet} \partial_\lambda \mathrm{e}^{-\frac{1}{2} \|\lambda\|_{\mathscr{P}}^2} \big|_{w=0} \\
&= -4i u^{\mathrm{T}} \vartheta \mathscr{Q}_{2\bullet} \mathscr{P} \begin{bmatrix} 0 \\ u \end{bmatrix} \mathrm{e}^{-\frac{1}{2} \|u\|_{\mathscr{P}_{22}}^2} \\
&= -4i u^{\mathrm{T}} \mathbf{S}(\vartheta \mathscr{E}_{22}) u \mathrm{e}^{-\frac{1}{2} \|u\|_{\mathscr{P}_{22}}^2},
\end{aligned}
\tag{56}
$$

where use is made of the Hankelian from (44). By substituting (56) into (52), the Gateaux derivative of $\mathscr{L}$ along K in (50) takes the form $i \mathbf{E}_* [K, \sigma] = 4u^{\mathrm{T}} \mathbf{S}(\vartheta \mathscr{E}_{22}) u \mathrm{e}^{-\frac{1}{2} \|u\|_{\mathscr{P}_{22}}^2} \mathrm{Re}\,\alpha$. Therefore, the fulfillment of (45) makes this derivative vanish for any $\alpha \in \mathbb{C}$ and $u \in \mathbb{R}^\nu$. Similar arguments (see also (49)) lead to the following Gateaux derivative of $\mathscr{L}$ along M in (51):

$$
\begin{aligned}
& 8 \mathrm{Im} \mathbf{E}_* \left(\mathscr{X}^{\mathrm{T}} \mathscr{2} \Theta \begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}} \Pi \end{bmatrix} \Omega \Pi^{\mathrm{T}} \mathrm{Re}(\beta \mathsf{W}_u) \right) \\
& - \mathrm{Re} \mathbf{E}_* \left[\mathscr{X}^{\mathrm{T}} \begin{bmatrix} 2 N^{\mathrm{T}} \\ N_1^{\mathrm{T}} \Pi \end{bmatrix} \Omega \Pi^{\mathrm{T}} \mathrm{Re}(\beta \mathsf{W}_u), \sigma \right] \\
=& 8 \mathrm{Im} \left\langle \mathscr{2} \Theta \begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}} \Pi \end{bmatrix} \overline{\Omega} \Pi^{\mathrm{T}}, \mathbf{E}_* (\mathscr{X} \mathrm{Re}(\beta^{\mathrm{T}} \mathsf{W}_u)) \right\rangle \\
& + \mathrm{Re} \left\langle \begin{bmatrix} 2 N^{\mathrm{T}} \\ N_1^{\mathrm{T}} \Pi \end{bmatrix} \overline{\Omega} \Pi^{\mathrm{T}}, \mathbf{E}_* [\sigma, \mathscr{X} \mathrm{Re}(\beta^{\mathrm{T}} \mathsf{W}_u)] \right\rangle \\
=& -8 \mathrm{e}^{-\frac{1}{2} \|u\|_{\mathscr{P}_{22}}^2} \mathrm{Im} \left\langle \mathscr{2} \Theta \begin{bmatrix} N^{\mathrm{T}} \\ N_1^{\mathrm{T}} \Pi \end{bmatrix} \overline{\Omega} \Pi^{\mathrm{T}}, \left(\mathscr{P}_{\bullet 2} + \begin{bmatrix} 0 \\ i\vartheta \end{bmatrix} \right) u \mathrm{Im} \beta^{\mathrm{T}} \right\rangle \\
& + 4 \mathrm{e}^{-\frac{1}{2} \|u\|_{\mathscr{P}_{22}}^2} \mathrm{Re} \left\langle \begin{bmatrix} 2 N^{\mathrm{T}} \\ N_1^{\mathrm{T}} \Pi \end{bmatrix} \overline{\Omega} \Pi^{\mathrm{T}}, i(\Theta \mathscr{E}_{12} - \mathscr{E}_{21}^{\mathrm{T}} \vartheta) u \mathrm{Im} \beta^{\mathrm{T}} \right\rangle \\
=& 4 \mathrm{Im} \beta^{\mathrm{T}} \Pi (J b_1^{\mathrm{T}} \mathscr{E}_{22} - (C \mathscr{E}_{21}^{\mathrm{T}} + B^{\mathrm{T}} \mathscr{Q}_{12} + b_1^{\mathrm{T}} \mathscr{Q}_{22}) \vartheta) u \mathrm{e}^{-\frac{1}{2} \|u\|_{\mathscr{P}_{22}}^2}.
\end{aligned}
\tag{57}
$$

Here, $\mathbf{E}_*[\sigma, \mathsf{W}_u] = 0$ follows from (56) under the condition (45). Also, in view of (53)–(55), use is made of the relations

$$\mathbf{E}_*(\mathscr{X}\mathsf{W}_u) = -i\partial_\lambda \mathbf{E}_* \mathrm{e}^{i\lambda^\mathrm{T}\mathscr{X}}\big|_{w=0} - \begin{bmatrix} 0 \\ \vartheta u \end{bmatrix} \mathbf{E}_*\mathsf{W}_u$$

$$= \mathrm{e}^{-\frac{1}{2}\|u\|^2_{\mathscr{P}_{22}}} \left(i\mathscr{P}_{\bullet 2} - \begin{bmatrix} 0 \\ \vartheta \end{bmatrix} \right) u,$$

$$\mathbf{E}_*[\sigma, \mathscr{X}\mathsf{W}_u] = -i\partial_\lambda \mathbf{E}_*[\sigma, \mathrm{e}^{i\lambda^\mathrm{T}\mathscr{X}}]\big|_{w=0} - \begin{bmatrix} 0 \\ \vartheta u \end{bmatrix} \mathbf{E}_*[\sigma, \mathsf{W}_u]$$

$$= -4\partial_\lambda(\lambda^\mathrm{T}\boldsymbol{\Theta}\mathscr{Q}\partial_\lambda \mathbf{E}_* \mathrm{e}^{i\lambda^\mathrm{T}\mathscr{X}})\big|_{w=0}$$

$$= 4\partial_\lambda\left(\lambda^\mathrm{T}\mathbf{S}(\boldsymbol{\Theta}\mathscr{E})\lambda \mathrm{e}^{-\frac{1}{2}\|\lambda\|^2_{\mathscr{P}}}\right)\big|_{w=0}$$

$$= 4\mathrm{e}^{-\frac{1}{2}\|u\|^2_{\mathscr{P}_{22}}}(\boldsymbol{\Theta}\mathscr{E}_{12} - \mathscr{E}_{21}^\mathrm{T}\vartheta)u.$$

The fulfillment of (45) and (46) makes the Gateaux derivative (57) of $\mathscr{Z}$ along the Weyl variation M in (51) also vanish for any $\beta \in \mathbb{C}^p$ and $u \in \mathbb{R}^\nu$. Therefore, (45) and (46) indeed imply the stationarity with respect to the arbitrary Weyl variations (50) and (51). ∎

Note that the proof of Theorem 2 (which employs the Gaussian averaging of quasi-polynomials of system variables, such as $\mathscr{X}\mathrm{e}^{i\lambda^\mathrm{T}\mathscr{X}}$) is closely related to the integro-differential identities [29] for expectations of Weyl quantization integrals over Gaussian states.

VIII. CONCLUSION

We have applied the transverse Hamiltonian variational method [30] to the mean square optimal CQF problem for a linear observer cascaded with a linear quantum plant. It has been shown that if such an observer is a stationary point of the problem among linear observers, then it also satisfies the first-order necessary conditions of optimality with respect to a wider class of Weyl variations of the energy operators of the observer. In this sense, linear observers are locally sufficient for linear quantum plants as far as the mean square performance criteria are concerned. Similar ideas can be developed for the CQLQG control problem [18] and an optimal control theory for classical port-Hamiltonian systems [26] and their stochastic versions.

REFERENCES

[1] B.D.O.Anderson, and J.B.Moore, *Optimal Control: Linear Quadratic Methods*, Prentice Hall, London, 1989.
[2] C.D.Cushen, and R.L.Hudson, A quantum-mechanical central limit theorem, *J. Appl. Prob.*, vol. 8, no. 3, 1971, pp. 454–469.
[3] D.Dong, and I.R.Petersen, Quantum control theory and applications: a survey, *IET Contr. Theor. Appl.*, vol. 4, no. 12, 2010, pp. 2651–2671.
[4] S.C.Edwards, and V.P.Belavkin, Optimal quantum filtering and quantum feedback control, arXiv:quant-ph/0506018v2, August 1, 2005.
[5] G.B.Folland, *Harmonic Analysis in Phase Space*, Princeton University Press, Princeton, 1989.
[6] C.W.Gardiner, and P.Zoller, *Quantum Noise*, 3rd Ed., Springer, Berlin, 2004.
[7] V.Gorini, A.Kossakowski, and E.C.G.Sudarshan, Completely positive dynamical semigroups of *N*-level systems, *J. Math. Phys.*, vol. 17, no. 5, 1976, pp. 821–825.
[8] J.Gough, and M.R.James, Quantum feedback networks: Hamiltonian formulation, *Commun. Math. Phys.*, vol. 287, 2009, pp. 1109–1132.
[9] A.S.Holevo, Quantum stochastic calculus, *J. Math. Sci.*, vol. 56, no. 5, 1991, pp. 2609–2624.
[10] R.L.Hudson, and K.R.Parthasarathy, Quantum Ito's Formula and Stochastic Evolutions, *Commun. Math. Phys.*, vol. 93, 1984, pp. 301–323.
[11] K.Jacobs, and P.L.Knight, Linear quantum trajectories: applications to continuous projection measurements, *Phys. Rev. A*, vol. 57, no. 4, 1998, pp. 2301–2310.
[12] M.R.James, H.I.Nurdin, and I.R.Petersen, H^∞ control of linear quantum stochastic systems, *IEEE Trans. Automat. Contr.*, vol. 53, no. 8, 2008, pp. 1787–1803.
[13] I.Karatzas, and S.E.Shreve, *Brownian Motion and Stochastic Calculus*, 2nd Ed., Springer-Verlag, New York, 1991.
[14] H.Kwakernaak, and R.Sivan, *Linear Optimal Control Systems*, Wiley, New York, 1972.
[15] G.Lindblad, On the generators of quantum dynamical semigroups, *Commun. Math. Phys.*, vol. 48, 1976, pp. 119–130.
[16] E.Merzbacher, *Quantum Mechanics*, 3rd Ed., Wiley, New York, 1998.
[17] Z.Miao, and M.R.James, Quantum observer for linear quantum stochastic systems, Proc. 51st IEEE Conf. Decision Control, Maui, Hawaii, USA, December 10-13, 2012, pp. 1680–1684.
[18] H.I.Nurdin, M.R.James, and I.R.Petersen, Coherent quantum LQG control, *Automatica*, vol. 45, 2009, pp. 1837–1846.
[19] K.R.Parthasarathy, *An Introduction to Quantum Stochastic Calculus*, Birkhäuser, Basel, 1992.
[20] K.R.Parthasarathy, What is a Gaussian state? *Commun. Stoch. Anal.*, vol. 4, no. 2, 2010, pp. 143–160.
[21] K.R.Parthasarathy, and K.Schmidt, *Positive Definite Kernels, Continuous Tensor Products, and Central Limit Theorems of Probability Theory*, Springer-Verlag, Berlin, 1972.
[22] I.R.Petersen, Quantum linear systems theory, Proc. 19th Int. Symp. Math. Theor. Networks Syst., Budapest, Hungary, July 5–9, 2010, pp. 2173–2184.
[23] L.S.Pontryagin, V.G.Boltyanskii, R.V.Gamkrelidze, and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Wiley, New York, 1962.
[24] A.K.Sichani, I.G.Vladimirov, and I.R.Petersen, Robust mean square stability of open quantum stochastic systems with Hamiltonian perturbations in a Weyl quantization form, Proc. Australian Control Conference Canberra, 17-18 November 2014, pp. 83–88.
[25] A.K.Sichani, I.G.Vladimirov, and I.R.Petersen, A gradient descent approach to optimal coherent quantum LQG controller design, *accepted at American Control Conference, 2015*, preprint: arXiv:1502.00274, 1 February 2015.
[26] A. Van der Schaft, Port-Hamiltonian systems: an introductory survey, Proc. Internat. Congress of Math., Madrid, Spain, 2006.
[27] I.G.Vladimirov, and I.R.Petersen, A quasi-separation principle and Newton-like scheme for coherent quantum LQG control, *Syst. Contr. Lett.*, vol. 62, no. 7, 2013, pp. 550–559.
[28] I.G.Vladimirov, and I.R.Petersen, Coherent quantum filtering for physically realizable linear quantum plants, Proc. European Control Conference, IEEE, Zurich, Switzerland, 17-19 July 2013, pp. 2717–2723.
[29] I.G.Vladimirov, A quantum mechanical version of Price's theorem for Gaussian states, Proc. Australian Control Conference, AUCC 2014, IEEE, Canberra, Australia, 17-18 November 2014, pp. 118–123, preprint: arXiv:1409.4377 [quant-ph], 15 September 2014.
[30] I.G.Vladimirov, A transverse Hamiltonian variational technique for open quantum stochastic systems and its application to coherent quantum control, *submitted*, preprint: arXiv:1506.04737 [quant-ph], 15 June 2015.
[31] R.M.Wilcox, Exponential operators and parameter differentiation in quantum physics, *J. Math. Phys.*, vol. 8, no. 4, 1967, pp. 962–982.